Ya Ju Fan

Center for Applied Scientific Computing
Lawrence Livermore National Laboratory
Livermore, CA 94551

Phone: (925) 423-4438
E-mail: fan4@llnl.gov

RESEARCH INTERESTS

- Statistical data analysis
- Mathematical models
- Pattern recognition and data mining
- Machine learning algorithms

EDUCATION

Ph.D. Industrial & Systems Engineering, Rutgers University, New Brunswick, NJ, 2010
Thesis: *Optimization models and algorithms for sample-preserved classification and clustering*

M.S. Industrial & Systems Engineering, University of Wisconsin – Madison, 2005
Major in Decision Science/Operations Research
Thesis: *Algorithms for solving the nonlinear programming problem of likelihood basis pursuit model*

B.B.A. Production & Operations Management, Fu Jen Catholic University, Taiwan, 2001

WORK/RESEARCH EXPERIENCE

Postdoctoral Researcher, CASC, LLNL, November 2010 – Present

- Studying and actively developing algorithms for anomaly /outlier detection for multiple-sensor streaming data
- Applied clustering algorithms for identifying time series motifs in wind generation data
- Developed integer solution in bilinear programming algorithm for solving a local region covering problem
- Compared dimension reduction and statistical feature selection methods for practical applications
- Implemented variations of nonlinear dimension reduction techniques for large scale scientific data analysis

Research Assistant, Rutgers University, July 2006 – June 2010

- Developed sample-preserved clustering algorithms for multivariate data
- Developed classification algorithms for multivariate time series analysis

Teaching Assistant, Rutgers University, 2008 – 2009

Database Developer, Vlife Technology Co. Ltd., Taiwan, Oct 2005 – Feb 2006

- Designed relational Oracle databases for a life insurance company
- Carefully understood user's requirements, implemented the business processes into PL/SQL language, and used Java and XML for user-interface

Research Assistant, UW-Madison, Department of Industrial Engineering, Sept 2002 – Dec 2004

- Solved the Likelihood Basis Pursuit (LBP) Model for Probability Estimation
- Large-Scale Nonlinear Regression Analysis on Real-world Medical Data Using LBP
- A Study of Polyhedral Analysis of Lot-sizing Problem

Teaching Assistant, UW-Madison, 2003 – 2004

HONORS & AWARDS

- Pierskalla Best Paper Award, INFORMS Annual Meeting, Washington, D.C., 2008
- Transportation Coordinating Council/Federal Transit Administration (TCC/FTA) Fellowship, 2007-2008
- Kuhl Memorial Engineering Fellowship, Rutgers Graduate School, 2006

PUBLICATIONS

C. Kamath and **Y. J. Fan**, "Using Data Mining Techniques to Enable Integration of Wind Energy on the Power Grid". *Statistical Analysis & Data Mining*. Volume 5, Issue 5, October 2012, pp 410-427.

C. Kamath and **Y. J. Fan**, "Finding Motifs in Wind Generation Time Series Data",*11th Conference on Machine Learning and Applications ICMLA 2012*. December 2012.

Y. J. Fan and C. Kamath. "On the Selection of Dimension Reduction Techniques for Scientific Applications". *Annals of Information Systems*. To Appear. August 2012.

Y. J. Fan and W. A. Chaovalitwongse. "Optimizing Feature Selection to Improve Medical Diagnosis". *Annals of Operations Research*, 174(1), pp. 169-183, February, 2010.

W. Chaovalitwongse, R. S. Pottenger, **Y. J. Fan**, S. Wang and L. D. Iasemidis. "Pattern-Based and Network-Based Classification Techniques for Multichannel Medical Data Signals to Improve Brain Diagnosis". Submitted to *Data Mining and Knowledge Discovery*, July 2009.

Y. J. Fan, W. A. Chaovalitwongse, C. Liu, R. C. Sachdeo, L. D. Iasemidis and P. M. Pardalos. "Optimisation and data mining techniques for the screening of epileptic patients". *International Journal of Bioinformatics Research and Applications*, 5(2): 187-196, March 2009.

W. Chaovalitwongse, **Y. J. Fan** and R.C. Sachdeo. "Novel Optimization Models for Abnormal Brain Activity Classification". *Operations Research*, 56(6): 1450-1460, December, 2008.

W. Chaovalitwongse, **Y. J. Fan** and R. C. Sachdeo. "On the Time Series K-Nearest Neighbor Classification of Abnormal Brain Activity". *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 37(6): 1005-1016, November, 2007.

W. Chaovalitwongse, **Y. J. Fan** and R.C. Sachdeo. "Support Feature Machine for Classification of Abnormal Brain Activity". The Thirteenth ACM SIGKDD International Conference On Knowledge Discovery and Data Mining (SIGKDD 2007), pp. 113-122.

Y. J. Fan, C. Iyigun and W. Chaovalitwongse. "Recent Advances in Optimization Models for Data Mining: Clustering and Classification". *CRM Proceedings & Lecture Notes of the American Mathematical Society (AMS)*, 45: 67-93, 2008.

Y. J. Fan and W. Chaovalitwongse. "Deterministic and Probabilistic Optimization Models for Data Classification". In C.A. Floudas and P.M. Pardalos (Eds.), *Encyclopedia of Optimization 2009*, Vol. II. Springer, New York, pp. 694-702.